

## SYNTHETIC BIOLOGY

## Accelerated evolution of a minimal 63–amino acid dual transcription factor

Andreas K. Brödel<sup>1</sup>, Rui Rodrigues<sup>2</sup>, Alfonso Jaramillo<sup>2,3,4</sup>, Mark Isalan<sup>1\*</sup>

Transcription factors control gene expression in all life. This raises the question of what is the smallest protein that can support such activity. In nature, Cro from bacteriophage  $\lambda$  is one of the smallest known repressors (66 amino acids), and activators are typically much larger (e.g.,  $\lambda$  cI, 237 amino acids). Previous efforts to engineer a minimal activator from  $\lambda$  Cro resulted in no activity in vivo in cells. In this study, we show that directed evolution results in a new Cro activator-repressor that functions as efficiently as  $\lambda$  cI in vivo. To achieve this, we develop phagemid-assisted continuous evolution (PACEmid). We find that a peptide as small as 63 amino acids functions efficiently as an activator and/or repressor. To our knowledge, this is the smallest protein activator that enables polymerase recruitment, highlighting the capacity of transcription factors to evolve from very short peptide sequences.

Copyright © 2020  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
License 4.0 (CC BY).

## INTRODUCTION

DNA binding proteins that regulate transcription initiation and control gene expression are called transcription factors (TFs). The question of, “What is the smallest peptide that can function as a TF?” is a fundamental one, with broad implications for the evolution of gene regulation (1–3). However, the potential for a peptide to be a minimal TF depends on its function: whether it is an activator, a repressor, or has dual activity (table S1).

In bacteria, TFs usually recruit or block RNA polymerase to activate or repress genes of interest, respectively. Generally, repression is more straightforward to achieve because it only requires a DNA binding protein to occlude key motifs in a promoter or to block transcription elongation by “roadblock” (4). By contrast, activator TFs have to strike a balance between DNA binding, RNA polymerase recruitment, and RNA polymerase release to initiate transcription efficiently (5). Consequently, activators and dual TFs might need to be larger and more complex than repressors (table S1). These include well-studied examples such as  $\lambda$  cI (237 amino acids) from bacteriophage  $\lambda$ 's genetic switch (fig. S1A) (6).

In this study, we set out to test the minimal size limits for TFs, and in particular whether activators or dual TFs could be made as small as the smallest known repressors. Viruses have some of the smallest functional TFs, and phage  $\lambda$  Cro protein is one of the smallest TF repressors characterized to date, containing only 66 amino acids (7, 8). Cro controls the viral life cycle along with  $\lambda$  cI, and together, they function naturally as a toggle switch (fig. S1B) (6). Other examples of very small repressors are Cro from *Enterobacteria* phage P22 (61 amino acids) (9) and CopG (45 amino acids) (10).

It has been previously shown that Cro might potentially be converted into an activator by transferring the activating surface patch of  $\lambda$  cI onto the surface of Cro (11). This can be done as molecular models of  $\lambda$  cI and Cro suggest that  $\alpha$  helices two and three of each protein lie in nearly identical positions. As a result,  $\alpha$  helices of both TFs are positioned to interact with the same part of RNA polymerase,

and Cro might become an activator if  $\alpha$  helix two were suitably engineered. However, a rationally engineered Cro variant had only a trace of detectable activity in vitro and none in vivo, in *Escherichia coli*, because of its low affinity for  $\lambda$  operators (11). Cro is significantly smaller than common TFs [e.g., LacI, 360 amino acids; TetR, 207 amino acids (table S1)], making it an ideal scaffold for developing a new set of small or minimal activators or dual TFs (fig. S1B). In addition, a set of Cro activators would complement the  $\lambda$  cI synthetic biology toolbox (12) for gene circuit engineering in bacterial cells (13, 14), because cI and Cro function on related operators.

## RESULTS

## M13 phagemid-assisted continuous evolution

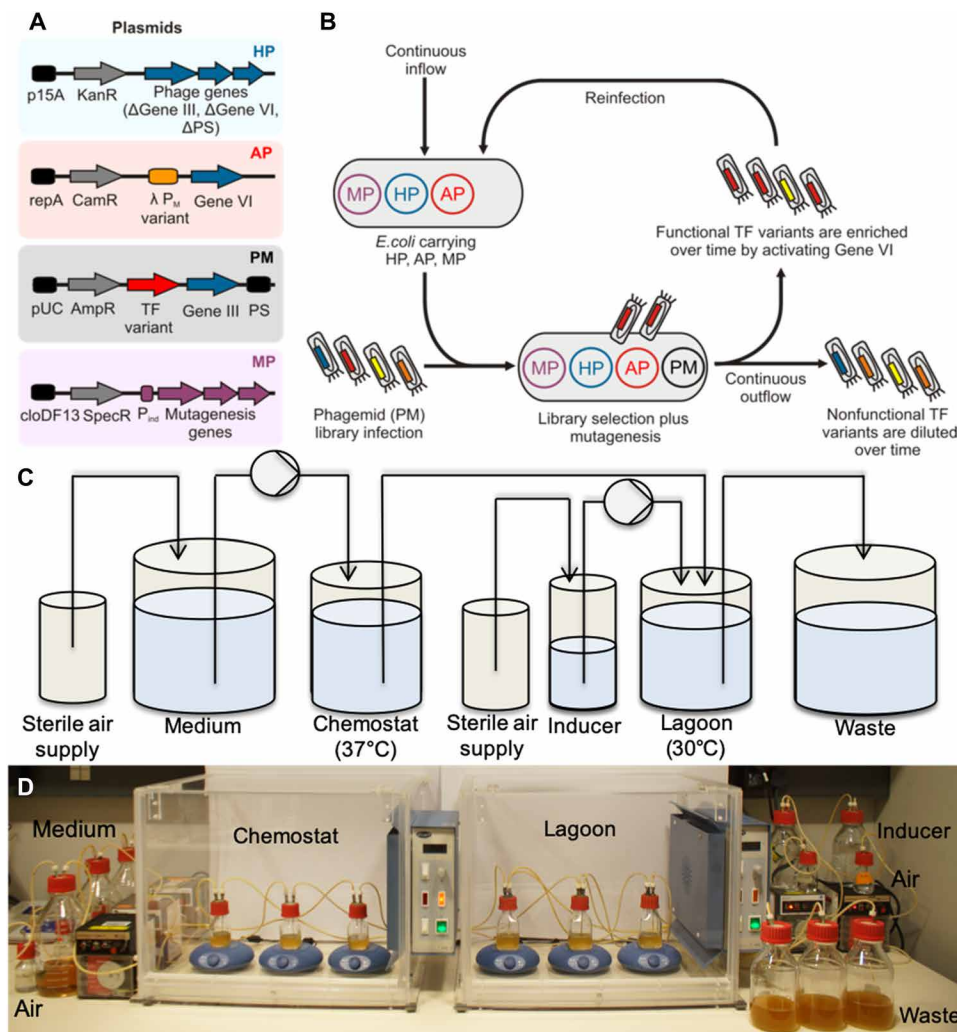
To obtain a set of small transcriptional activators derived from  $\lambda$  Cro, we converted our phagemid-based batch selection system (12, 15) for accelerated continuous evolution in a manner similar to phage-assisted continuous evolution (PACE) (16, 17). One difference between our system and classic PACE is that we use phagemids rather than phage to facilitate screening large libraries. The smaller size of the phagemid enables the construction of gene libraries with a much higher number of variants than standard phage. We also prevent bystander mutations in phage genes by placing them on helper plasmids (HP) and accessory plasmids (AP), respectively (Fig. 1A). These plasmids are continuously replenished within fresh uninfected cells (Fig. 1B). The resulting phagemid evolution system, which we call phagemid-assisted continuous evolution (PACEmid), is based on conditional M13 bacteriophage replication. The selection process takes place inside *E. coli* cells by linking the evolving TF activity to restoring essential phage Gene VI expression (deleted from the HP). In this way, a TF with novel properties can be selected after several cycles of reinfection (Fig. 1B), and the process can be automated (Fig. 1, C and D).

For continuous selection, we found it essential to tune the basal Gene VI expression rate to produce sufficient amounts of phage in the absence of an active TF, reducing the chances of phage loss in the lagoon. We carried out model selections with cI<sub>opt</sub> [a  $\lambda$  cI optimized mutant with a strong activation region (18)] and showed that selection stringency and rate can be tuned by changing the copy number (19) of the AP (fig. S2). Furthermore, enrichment of cI<sub>opt</sub> was more efficient in continuous mode than in batch mode under

<sup>1</sup>Department of Life Sciences, Imperial College London, London SW7 2AZ, UK.

<sup>2</sup>Warwick Integrative Synthetic Biology Centre and School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK. <sup>3</sup>CNRS-UMR8030, Laboratoire ISSB and Université Paris-Saclay and Université d'Évry and CEA, DRF, IG, Genoscope, Évry 91000, France. <sup>4</sup>Institute for Integrative Systems Biology (I2SysBio), University of Valencia-CSIC, 46980 Paterna, Spain.

\*Corresponding author. Email: m.isalan@imperial.ac.uk



**Fig. 1. PACEmid of TFs.** (A) Plasmids: HP, helper phage to provide all phage genes except for gIII and gVI; AP, AP to provide conditional Gene VI expression to enable selection of a successful evolving TF variant; PM, phagemid containing an evolving TF variant and gIII; MP, A chemically inducible mutagenesis plasmid (MP) can, optionally, be used in addition to combinatorial library selection. (B) Continuous selection flow diagram: Host cells carrying the HP, AP, and, optionally, an MP continuously flow from a chemostat into a lagoon where they get infected with M13 phage. Only an active TF induces Gene VI expression to complete the phage life cycle, thus enriching this library variant; nonfunctional TF variants are diluted over time. Diversification of the target gene is obtained by combinatorial libraries and/or random mutagenesis. (C) Flow chart of the PACEmid continuous evolution system. *E. coli* cells (containing HP, AP, and MP) are cultured in the late log phase (chemostat, 37°C) and flow through a lagoon (30°C) containing the evolving phagemid (PM). (D) Photo of bioreactor setup showing three independent experiments performed in parallel (Photo credit: M. Mielcarek, Imperial College London).

the same selection pressure (fig. S3), confirming the advantages of continuous selection (16, 17) for accelerated evolution.

We next implemented a mutagenesis device to expand the mutation spectrum beyond combinatorial libraries. On ColE1-derived plasmids, such as our phagemid, the leading-strand replication pol I is gradually replaced by pol III over at least 1.3 kb downstream of the origin of replication (20). We therefore characterized the efficiency of three mutagenesis cassettes carrying error-prone pol I [EP pol I (21)] or error-prone pol III variants [MP4 and MP6 (22)] under the inducible promoters P<sub>BAD</sub> and P<sub>Lac</sub> (23). Mutation rates were then analyzed by a  $\beta$ -lactamase stop codon reversion assay (fig. S4A), a rifampicin resistance assay (fig. S4, B and C), and by monitoring loss of red fluorescent protein (RFP) function on M13 phagemids after three rounds of batch evolution (fig. S4, D and E). The use of the MP6 (22) cassette under the isopropyl  $\beta$ -D-1-thiogalactopyranoside

(IPTG)-inducible promoter P<sub>Lac</sub> (P<sub>Lac</sub>-MP6-SpecR) led to the highest mutation rates, and this was chosen for downstream applications (fig. S4F). When induced with arabinose, the original MP6 leads to a 322,000-fold increase in mutation rate over that of wild-type (WT) *E. coli*. This results in an average of  $6.2 \times 10^{-6}$  substitutions per base pair (bp) per generation (22).

To validate the use of the adapted MP6 (22) mutator cassette for directed evolution, we first evolved an improved orthogonal  $\lambda$  cI TF [cI<sub>4A5T6T,P</sub>; formerly the least active member of our cI toolkit (12)] (fig. S5A). The resulting cI variant had a Met-to-Thr mutation at position 42, leading to an improved dual activation-repression of green fluorescent protein (GFP) and mCherry (fig. S5B). The evolved cI<sub>4A5T6T,P</sub> (T42) up-regulated GFP expression 5.4-fold and led to a 91% repression of mCherry (fig. S5, C and D) (full sequences in the Supplementary Materials). In comparison, expression of the parental

$\text{cI}_{4\text{A5T6T,P (M42)}}$  only displayed a 4.3-fold activation and a 58% repression. Notably, position 42 was not randomized in our previously constructed  $\lambda$  cI library and, thus, confirmed the potential of extra mutagenesis for improved protein activities.

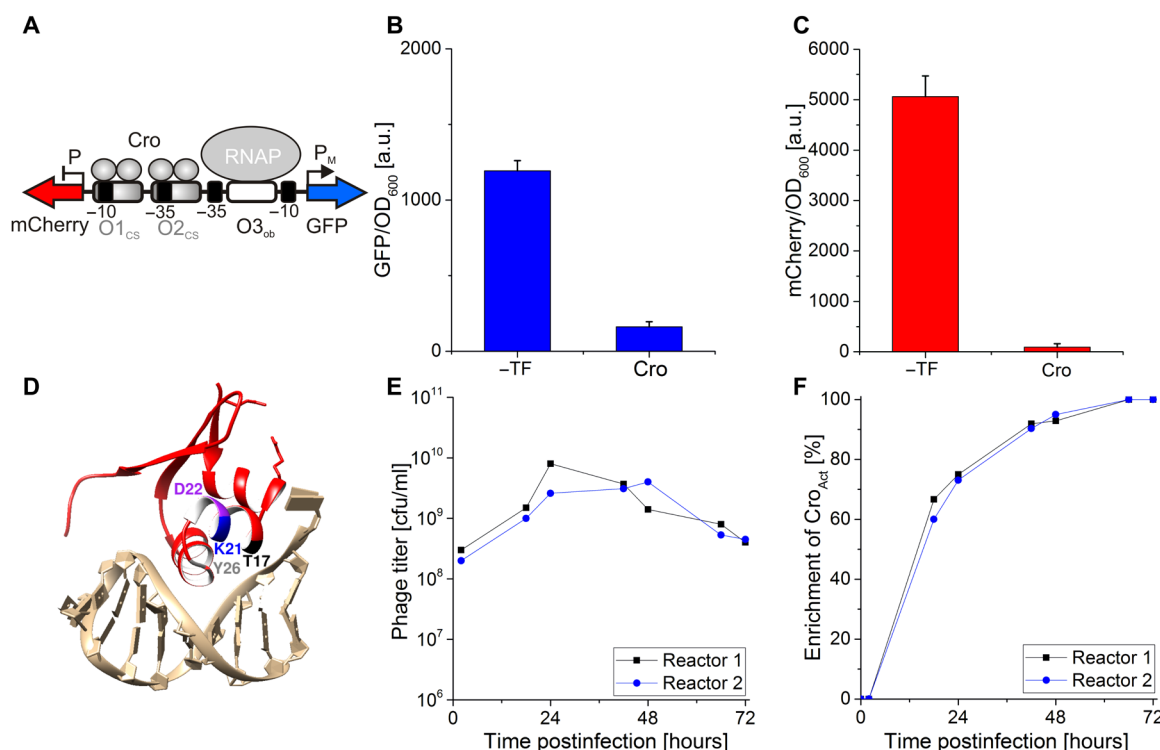
### Directed evolution of Cro activators

To see whether the optimized phagemid selection system was powerful enough to select a small Cro activator that could function in *E. coli*, we first constructed a bidirectional promoter ( $P_{\text{CS}}/P_{\text{M,CS}}$ ), with three operator sites ( $O_1$ ,  $O_2$ ,  $O_3$ ) (Fig. 2A).  $O_1$  and  $O_2$  consisted of the  $\lambda$  operator consensus sequence (CS), with the highest reported affinity for  $\lambda$  Cro (24, 25). As Cro binds with a very high affinity to WT  $O_3$  (24), a mutated  $O_3$  site was used to reduce potential auto-repression of the  $P_{\text{M,CS}}$  promoter by  $\lambda$  Cro. Basal promoter activity of  $P_{\text{CS}}/P_{\text{M,CS}}$ , as well as the effect of Cro on expression levels, was characterized by GFP and mCherry production and showed that WT Cro repressed in both promoter directions (Fig. 2, B and C).

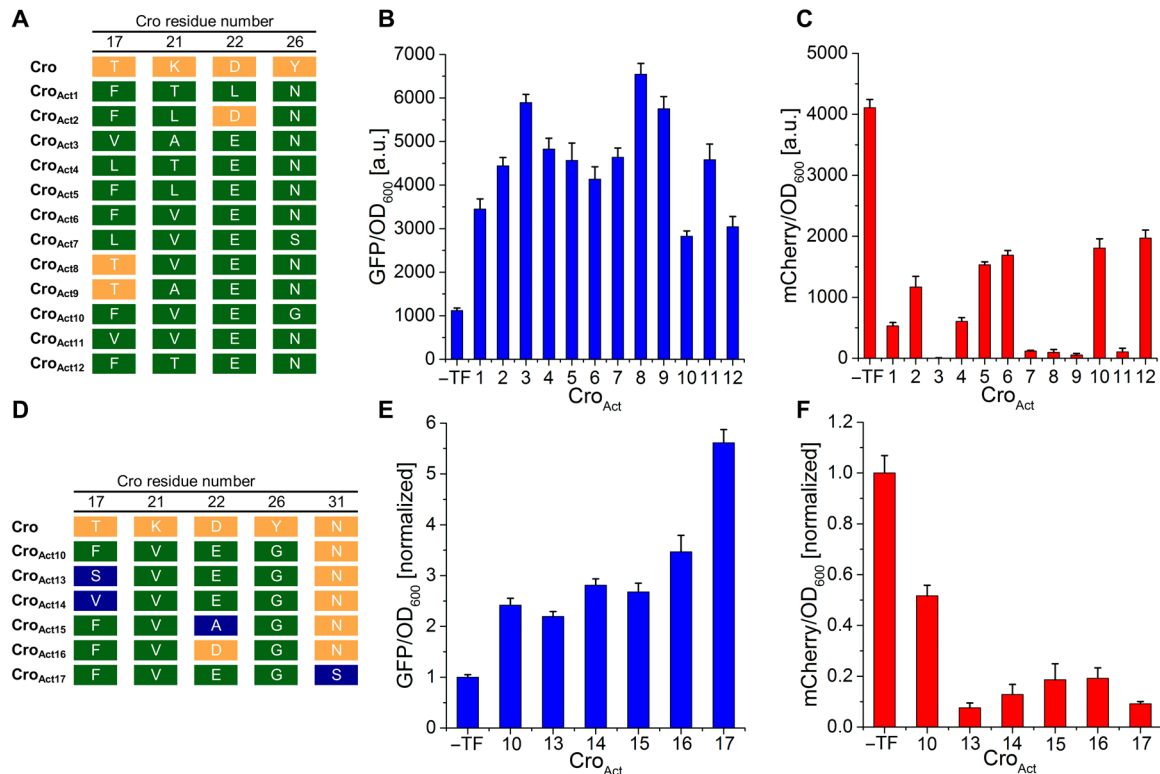
To search for activators, a combinatorial library of Cro variants was constructed by randomizing four amino acids in  $\alpha$  helix two and three (11) (Fig. 2D). Three residues (T17, K21, and D22) were in the potential activation patch, whereas the fourth residue (Y26) was upstream of the  $\alpha$  helix necessary for DNA binding. We then constructed an AP with Gene VI under the engineered  $P_{\text{M,CS}}$  promoter, on the pSC101 vector for high selection stringency. The

combinatorial Cro library was selected against this AP for 3 days in continuous culture, leading to a reproducible enrichment of Cro activators in two separate bioreactors (Fig. 2, E and F). The selected activators had at least three amino acid substitutions at the randomized positions over WT Cro (Fig. 3A). Notably, 10 of the 12 selected Cro activators contained an asparagine N at position 26. The importance of N26 for activity was confirmed with site-directed mutagenesis (fig. S6, A and B). The activity of the selected Cro activators ( $\text{Cro}_{\text{Act}}$ ) was then analyzed with the reporter assay (Fig. 2A).  $\text{Cro}_{\text{Act}}$  variants up-regulated GFP expression 2.3- to 5.8-fold and repressed mCherry between 52 and 100% (fig. 3, B and C). In comparison,  $\lambda$  cI expression resulted in a 6.7-fold activation and a 100% repression of the bidirectional  $P_{\text{R}}/P_{\text{RM}}$  promoter (fig. S6, C and D). The most active variants possessed the amino acids V17, A21, E22, and N26 ( $\text{Cro}_{\text{Act}3}$ ), and T17, V21, E22, and N26 ( $\text{Cro}_{\text{Act}8}$ ) at the randomized positions; both were strong activators and inhibitors of the dual promoter in Fig. 2A.

Next, we explored a wider mutation space using the mutagenesis cassette ( $P_{\text{Lac}}\text{-MP6-SpecR}$ ). To achieve this, we applied continuous directed evolution for 4 days, using an optimized pLITMUS\* vector backbone (fig. S7, A to C) and starting with the least-active variant  $\text{Cro}_{\text{Act}10}$  (Fig. 3A and fig. S7D). Directed evolution was performed under a medium selection pressure (medium copy number plasmid pJPC12) compared to library selections (low copy pSC101; strong



**Fig. 2. Directed evolution of Cro dual TF activator-repressors.** (A) Bidirectional promoter designed to activate GFP and repress mCherry. Operators:  $O_1$  and  $O_2$  sites recruit Cro;  $O_3$  is designed to have weaker Cro binding. RNAP, RNA polymerase. (B) Despite obliterating the operator  $O_3$ , expression of Cro still results in partial repression of GFP under the  $P_{\text{M,CS}}$  promoter due to some affinity to the operator 3. (C) Cro binds with high affinity to the operators  $O_1$  and  $O_2$  as a dimer, resulting in a strong repression of mCherry. GFP and mCherry expression was normalized to OD<sub>600</sub> (optical density at 600 nm), and data were obtained from four biological replicates. Error bars show 1 SD. (D) Structural model of WT Cro binding DNA, showing key residues (T17, K21, D22, and Y26) randomized in a combinatorial library (Protein Data Bank ID: 6CRO). (E and F) Time course of the phage titer and enrichment of Cro activators during selection in continuous culture in two bioreactors. Samples were taken from the outflow of each bioreactor, and phage supernatants were analyzed by infecting TG1 cells carrying the reporter plasmid. Cro activators up-regulate GFP and repress mCherry, which can be monitored by plate analysis (enrichment of Cro activators). cfu, colony-forming units. a.u., arbitrary units.



**Fig. 3. Analysis of Cro dual TF activator-repressors.** (A) Sequencing results of 12 selected dual Cro activators. WT amino acids are highlighted in orange, and non-WT amino acids in green. (B and C) Activation and repression of the bidirectional promoter  $P_{CS}/P_{MCS}$  by the selected Cro variants. GFP and mCherry expression was normalized to OD<sub>600</sub>. (D) Sequencing results from continuous directed evolution using the mutagenesis plasmid on least-active variant Cro<sub>Act10</sub>. (E and F) Activation and repression of the bidirectional promoter using the variants in (D). All data represent the average of four biological replicates, and error bars correspond to the SD between the measurements.

selection pressure). This resulted in higher phage production rates and, thus, an increased number of Cro variants. Five additional Cro activators were, thus, evolved with single amino acid changes in the polymerase interaction site (Cro<sub>Act13</sub> to Cro<sub>Act16</sub>) or DNA binding  $\alpha$  helix (Cro<sub>Act17</sub>) (Fig. 3, D to F). Notably, a new amino acid change at position 31, N31S in the DNA binding  $\alpha$  helix of Cro<sub>Act17</sub>, had a strong impact on the TF activity. To summarize, we obtained a set of 17 small Cro activators with a broad range of activities.

### Gene network engineering based on Cro activators

To test the potential for gene network engineering with the evolved Cro activators, we selected Cro<sub>Act3</sub> (5.3-fold activation, 100% repression) for use in combination with our set of orthogonal cI variants (12). Successful combination would allow the construction of a wide range of gene networks that integrate multiple inputs for activation and/or repression, based on variants of commonly used  $\lambda$  promoters. First, we verified the lack of cross-reactivity of the Cro<sub>Act3</sub> to any of the synthetic promoters of the orthogonal cI toolkit (fig. S8). Having found no unwanted cross-reactivity, we built two different gene circuits. In the first network, gradual addition of arabinose resulted in expression of Cro<sub>Act3</sub> and, thus, a concentration-dependent increase in GFP and a decrease in mCherry, as expected (Fig. 4, A and B). Note that arabinose addition generally leads to an induction of more cells rather than increased levels of all cells as the  $P_{BAD}$  promoter has an all-or-nothing response (26). In the second circuit, expression of the two orthogonal TFs Cro<sub>Act3</sub> and cI<sub>5G6G,P</sub>, linked to the inducers arab-

inose and 3OC6-HSL, resulted in a concentration-dependent increase in the reporters GFP and mCherry, as designed (Fig. 4, C and D). Therefore, Cro<sub>Act3</sub> can be used for gene circuit engineering, either alone or in combination with other orthogonal TFs, and can target unidirectional as well as bidirectional promoters in a concentration-dependent manner.

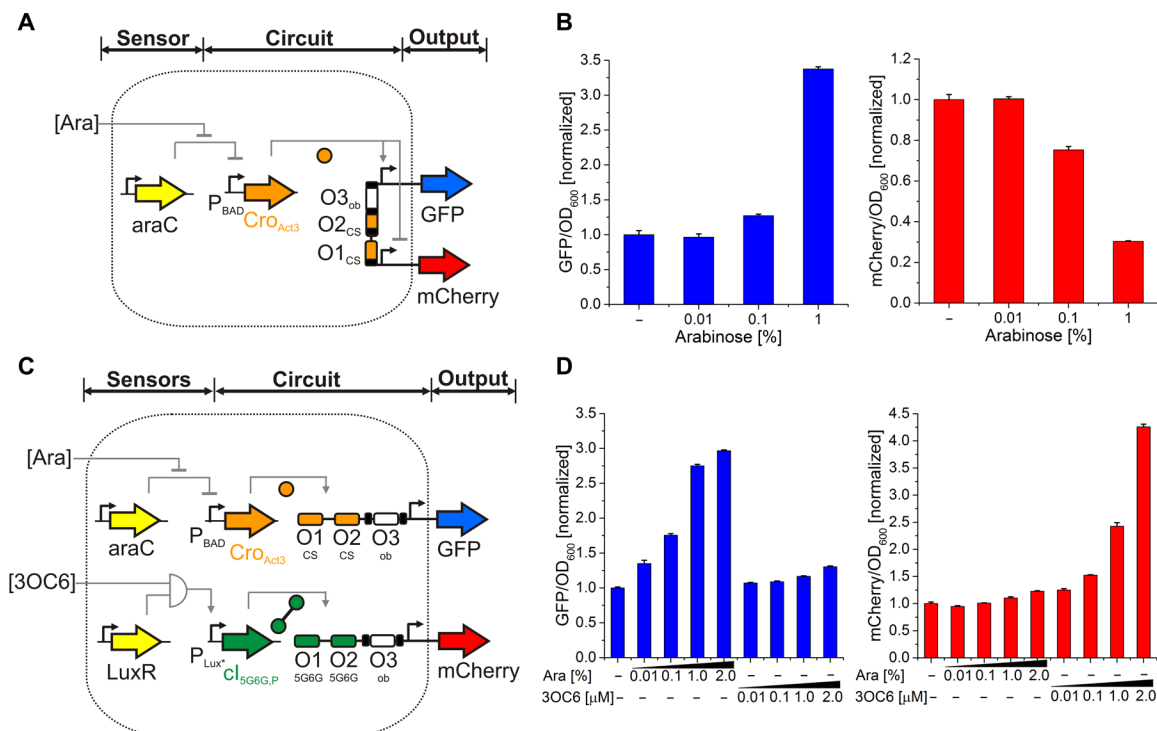
### Rational engineering of a minimal 63-amino acid dual TF

At 66 amino acids, Cro<sub>Act3</sub> already had a claim to being the smallest activator and dual TF that enables polymerase recruitment. However, we sought to push the boundaries of the smallest possible such TF. By identifying functional breakpoints in the expected structure (Protein Data Bank ID: 6CRO), we made targeted deletions to the C-terminal end of the TF and analyzed the variants using our reporter assay (Fig. 5A). Thus, we found that a minimal 63-amino acid protein (Cro<sub>Act3</sub><sub>63aa</sub>) is still capable of transcriptional activation or repression in *E. coli* (4.4-fold activation, 88% repression) (Fig. 5, B and C). To our knowledge, this makes it the smallest dual TF that enables polymerase recruitment.

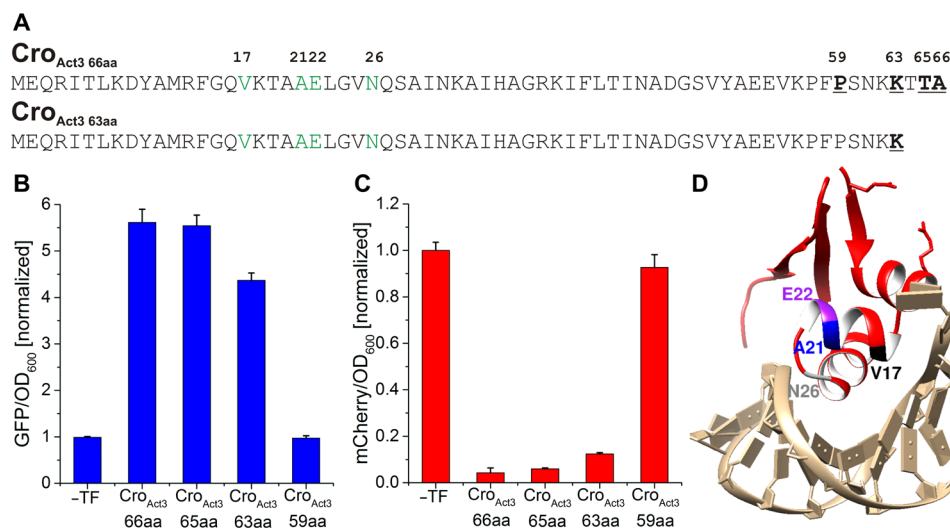
### DISCUSSION

We chose to evolve  $\lambda$  Cro because of its small size, its biological function as counterpart to  $\lambda$  cI, and its use in numerous synthetic biology projects (27–30). Furthermore, Cro's  $\alpha$  helix two is positioned to be able to interact with the same part of RNA polymerase that  $\lambda$  cI binds. The evolved Cro activators described in this study can easily





**Fig. 4. Synthetic gene circuits based on the evolved minimal activator  $\text{Cro}_{\text{Act3}}$ .** (A) Design of a one-input gene network on a bidirectional promoter. The arabinose (Ara)-inducible sensor induces  $\text{Cro}_{\text{Act3}}$  operating on a bidirectional promoter and two reporter genes. Operators:  $\text{O1}_{\text{CS}}$  and  $\text{O2}_{\text{CS}}$  sites (orange) recruit  $\text{Cro}_{\text{Act3}}$ ;  $\text{O3}_{\text{ob}}$  (white) is an obliterated site to circumvent TF binding. (B) Experimental data for the one-input system showing the concentration-dependent response of GFP and mCherry. (C) Design of a two-input gene network on two unidirectional promoters. Two sensors ( $\text{P}_{\text{BAD}}$  and  $\text{P}_{\text{Lux}^*}$ ) act on an integrating circuit with two orthogonal TFs ( $\text{Cro}_{\text{Act3}}$  and  $\text{Cl}_{5\text{G6G},\text{P}}$ ) operating on two unidirectional promoters and two reporter genes. Operators:  $\text{O1}_{\text{CS}}$  and  $\text{O2}_{\text{CS}}$  sites (orange) recruit  $\text{Cro}_{\text{Act3}}$ ;  $\text{O1}_{5\text{G6G}}$  and  $\text{O2}_{5\text{G6G}}$  sites (green) recruit  $\text{Cl}_{5\text{G6G},\text{P}}$ ;  $\text{O3}_{\text{ob}}$  (white) is an obliterated site to circumvent TF binding. (D) Experimental data for the two-input system illustrating the concentration-dependent response of GFP and mCherry to the inducers (Ara, 3OC6). All data represent the average of three biological replicates, and error bars correspond to the SD between the measurements.



**Fig. 5. Engineering of a minimal 63-amino acid dual TF.** (A) Functional breakpoints were identified in the sequence of  $\text{Cro}_{\text{Act3}}$ , including a potential structure-breaking  $\text{Pro}^{59}$ , a positive charge patch ending at  $\text{Lys}^{63}$ , and a potentially neutral C-terminal  $\text{Ala}^{66}$  (bold). The corresponding truncation mutants [59, 63, and 65 amino acids (aa)] were generated;  $\text{Cro}_{\text{Act3 } 63\text{aa}}$  is shown as an example. Activator mutations to WT  $\lambda$  Cro repressor are highlighted in green. (B and C) Activation and repression of the bidirectional promoter  $\text{P}_{\text{CS}}/\text{P}_{\text{MCS}}$  by truncated  $\text{Cro}_{\text{Act3}}$  variants. GFP and mCherry expression were normalized to  $\text{OD}_{600}$ ; four biological replicates; error bars represent SD between the measurements. Activation and repression were normalized to the basal expression of each promoter in the absence of any TF on the PM. (D) Model indicating key features of the minimal dual TF,  $\text{Cro}_{\text{Act3 } 63\text{aa}}$ .

be constructed by site-directed mutagenesis of WT  $\lambda$  Cro, providing a straightforward approach for users to implement these Cro variants into their synthetic biology projects (components are provided in figs. S9 to S11 and tables S2 to S6).

One basic scientific question that we address in this study is whether a small protein repressor can be converted into a transcriptional activator *in vivo*. Overall, we conclude that not only is this possible but also that even 63 amino acids are sufficient to carry out dual intracellular transcriptional repression and activation by polymerase recruitment ( $\sim 4.4$ -fold). It should be noted that TFs are distinct from DNA binding domains (DBDs). For example, certain zinc fingers can form functional DBDs, such as the 63-amino acid domain from GAGA factor [a single zinc-finger unit flanked by two basic rich regions (31)]. However, these interactions are not capable of transcription activation, and the full-length GAGA factor TF is 519 amino acids. A potential alternative to our engineered dual TF is the Arc protein (53 amino acids) from bacteriophage P22, which can act as a repressor or activator ( $\sim 2.8$ -fold) (32). However, activation does not occur here by recruiting a polymerase but via accelerating the rate at which the polymerase clears the promoter.

In this study, we used a powerful combination of rational design and directed evolution to obtain a set of Cro activators. An alternative strategy is to simply use directed evolution in the absence of a starting library. This would be very interesting as such an approach would lead to detailed information on specific evolution trajectories. In addition, this might give a time frame on how long it would take to walk through sequence space to obtain these beneficial mutations. Last, it would be interesting to further investigate whether our developed directed evolution platform can be easily transferred to evolve other small repressor peptides.

The small size of TFs raises interesting questions for the *de novo* evolution of DNA binding and polymerase-recruiting proteins from inert protein scaffolds. Cro belongs to the helix-turn-helix (HTH) superfamily (33), and this study implies that a relatively small amount of secondary structure, including three short  $\alpha$  helices and three  $\beta$  strands, is sufficient to make a compact scaffold that could support minimal gene activation. As long as there are many ways to reach similar folds, the subsequent number of mutations to make a TF may be more tractable for natural selection [ $\sim 11$  mutations for DNA binding (34) and  $\sim 2$  to 5 more for transcription activation]. We note that in yeast, a nine-amino acid peptide is sufficient for transactivation (35). Overall, this implies that short peptide scaffolds may have a greater capacity to evolve into DNA binding and RNA polymerase-recruiting proteins than previously thought.

## MATERIALS AND METHODS

### Strains and media

Standard DNA cloning was performed with chemically competent TOP10 cells (Invitrogen) and electrocompetent TG1 cells. Combinatorial library cloning was performed with NEB 5-alpha electrocompetent cells. Phage production was carried out with TOP10, S1030 (36), TG1, or BL21(DE3) cells. All phage-assisted infection assays and reporter assays were performed with TG1 cells. Genotypes of all strains are listed in table S2. Cells were grown in  $2\times$  TY medium [NaCl (5 g liter<sup>-1</sup>), yeast extract (10 g liter<sup>-1</sup>), tryptone (16 g liter<sup>-1</sup>)], M9 minimal medium [Na<sub>2</sub>HPO<sub>4</sub> (6.8 g liter<sup>-1</sup>), KH<sub>2</sub>PO<sub>4</sub> (3.0 g liter<sup>-1</sup>), NaCl (0.5 g liter<sup>-1</sup>), NH<sub>4</sub>Cl (1.0 g liter<sup>-1</sup>), 2 mM MgSO<sub>4</sub>, 100  $\mu$ M CaCl<sub>2</sub>, 0.2% (w/v) glucose, 1 mM thiamine-HCl], or Super Optimal

broth with Catabolite repression (SOC) medium (Sigma-Aldrich). Chloramphenicol (10 to 25  $\mu$ g ml<sup>-1</sup>), kanamycin (25 to 50  $\mu$ g ml<sup>-1</sup>), spectinomycin (25 to 50  $\mu$ g ml<sup>-1</sup>), ampicillin (50 to 100  $\mu$ g ml<sup>-1</sup>), tetracycline (5 to 10  $\mu$ g ml<sup>-1</sup>), and carbenicillin (10  $\mu$ g ml<sup>-1</sup>) were added where appropriate. IPTG, D-glucose, or L-arabinose was added to the media to induce or repress the promoter P<sub>Llac</sub> (23) or P<sub>BAD</sub>.

### Cloning and plasmid construction

Subcloning was carried out using Gibson Assembly (37). GFP (GenBank no. KM229386), RFP, and mCherry (UniProt no. X5DSL3) were used as reporters. The  $\lambda$  Cro regulatory protein (UniProt no. P03040) was used as TF scaffold, and the rpoN promoter (P<sub>rpoN</sub>) was used to express the evolving gene on the phagemid. The stronger activator cI<sub>opt</sub> (18) contains three amino acid changes in the  $\lambda$  cI gene (GenBank no. X00166) at positions 35 to 39 (SVADK to LVAYE). The change in copy number of APs (pSC101, pJPC12, and pJPC13) and the Cro<sub>Act3,Y26</sub> variant were obtained by site-directed mutagenesis. All mutagenesis plasmids [MP4 and MP6 (22); EP pol I and WT pol I (21)] were obtained from Addgene and recloned into a vector backbone carrying a spectinomycin resistance gene (SpecR) and a cloDF13 origin of replication to make it compatible with the other plasmids of the directed evolution system. The AraC-P<sub>BAD</sub> cassette on MP6-SpecR was replaced with the IPTG-inducible promoter P<sub>Llac</sub> (23) (P<sub>Llac</sub>-MP6-SpecR). For the construction of gene circuits, a modified version of the P<sub>Lux</sub> promoter (P<sub>Lux</sub>\*) (14) was used to reduce basal expression levels. A degradation tag (AANDENYALVA) was fused to cI<sub>5G6G,P</sub> at the C-terminal site to decrease basal expression in the absence of an inducer. Promoters, ribosomal binding sites, and terminators were ordered as oligonucleotides (Sigma-Aldrich) or were obtained from previous studies (38, 39). The P<sub>MCS</sub> promoter contained the mutated O3 sites TATAAATAGTGGTGATA (40) or ACAAACTTTCTTGTATA to bypass autorepression at high  $\lambda$  cI concentrations. Plasmids were purified using the QIAprep Spin Miniprep Kit or the HiSpeed Plasmid Maxi Kit (QIAGEN). Nucleotide sequences of all cloned constructs were confirmed by DNA sequencing (Eurofins Genomics). The DNA sequences of the synthetic promoters are listed in fig. S9. All plasmids and selected primer sequences are listed in tables S3 to S5. Maps for each class of plasmid are highlighted in fig. S10.

### Construction of a combinatorial Cro library

A combinatorial Cro library was cloned on the basis of forward and reverse primers carrying NNS codons (where S = G/C) at positions T17, K21, D22, and Y26, as described previously (table S6) (15). Primers were fused by polymerase chain reaction (PCR), and fragments were cloned into the linearized pLITMUS-P<sub>rpoN</sub>-Cro-P<sub>BBA<sub>J23106</sub></sub>-gIII vector by Gibson Assembly. Cells were transformed and plated on 24-cm<sup>2</sup> Nunc BioAssay Dishes (Thermo Fisher Scientific). Transformation efficiency was estimated by colony counting of plated serial dilutions. The next day, colonies were harvested and phagemid DNA was purified. Ten clones were sequenced to confirm diversity of the library (table S6). Molecular graphics of TFs were obtained with UCSF Chimera (41).

### Selection phage production

Selection phage production was performed in BL21(DE3) cells carrying HP- $\Delta$ PS- $\Delta$ gIII- $\Delta$ gVI and pJPC13- $\Delta$ PS-P<sub>T7</sub>-gVI or TOP10 cells containing HP- $\Delta$ PS- $\Delta$ gIII. Cells were made electrocompetent, phagemids were transformed, and cells were grown overnight at

30°C, 250 rpm (Stuart Shaking Incubator SI500), in 2× TY medium supplemented with kanamycin (12.5 µg ml<sup>-1</sup>), ampicillin (50 µg ml<sup>-1</sup>), and chloramphenicol (12.5 µg ml<sup>-1</sup>) where appropriate. For enrichment assays, plasmids carrying *cl<sub>opt</sub>* and RFP were mixed in a ratio of 10<sup>-6</sup> before transformation. IPTG (0.25 mM) was added to the BL21(DE3) culture after phagemid transformation to induce Gene VI expression. Samples were centrifuged for 10 min at 8000g, and supernatants were sterile filtered (0.22-µm pore size, Millex-GV). Phage concentration was analyzed by TG1 infection of diluted phage stocks and colony counting on ampicillin plates.

### Phagemid-assisted batch evolution

TG1 or S1030 cells carrying the helper phage HP-ΔPS-ΔgIII-ΔgVI, a Gene VI-based AP, and, optionally, a mutagenesis plasmid were grown on agar plates (M9 or LB) supplemented with appropriate antibiotics. For TG1-based evolution, starter cultures were inoculated in 2× TY with appropriate antibiotics and grown for 5 to 6 hours at 37°C until the OD<sub>600</sub> (optical density at 600 nm) reached 0.3 to 0.6. For S1030-based evolution, overnight cultures were inoculated from single colonies. The next day, selection cultures were prepared with a 100-fold dilution of the overnight cultures, and cells were grown for 3 to 4 hours at 37°C until the OD<sub>600</sub> reached 0.3 to 0.6.

Cultures were infected at a desired multiplicity of infection (MOI), and a chemical inducer was added where appropriate. Cell cultures were incubated for 20 hours at 30°C, 250 rpm (Stuart Shaking Incubator SI500). Overnight cultures were centrifuged for 10 min at 8000g, and the phage supernatant was used to start a new round of evolution. After each round, phage supernatants were diluted before infecting TG1 cells carrying an appropriate reporter plasmid. Infected cells were selected on ampicillin plates, and single colonies were grown overnight in 2× TY medium supplemented with ampicillin. Phagemid DNA was purified using the QIAprep Spin Miniprep Kit (QIAGEN) and analyzed by sequencing (Eurofins Genomics).

### Phagemid-assisted continuous evolution

Glass bottles, chemostats, and lagoons connected via biocompatible tubing (Cole-Parmer) were autoclaved and cooled down to room temperature. The autoclaved chemostats and lagoons were placed into two incubators (SID60, Stuart) on individual shakers (Topolino Mobil, IKA). Before each experiment, the bioreactor was equilibrated by pumping 2× TY supplemented with the appropriate antibiotics through the system at 20 ml hour<sup>-1</sup> (Pharmacia Biotech Pump P-1).

S1030 cells carrying the modified helper phage HP-ΔPS-ΔgIII-ΔgVI, an AP, and, optionally, a mutagenesis plasmid were grown on agar plates supplemented with appropriate antibiotics and 1% (w/v) D-glucose. The next day, 10-ml cultures were inoculated from single colonies, grown overnight at 37°C, 250 rpm, and stored at 4°C. Starter cultures were inoculated with a 100-fold dilution of the overnight culture and grown at 37°C, 250 rpm, until the OD<sub>600</sub> reached 0.3 to 0.6. Chemostats were filled with 25 ml of this starter culture, and cells were grown at 37°C with magnetic stir-bar agitation. The lagoon was filled with 40 ml of the starter culture, and cells were infected at a MOI of 4 and 1 mM IPTG was added where appropriate. The infected cells were grown at 30°C with magnetic stir-bar agitation. The flow rate of 2× TY supplemented with the appropriate antibiotics was set to 20 ml hour<sup>-1</sup> to provide continuous supply of media. IPTG (10 mM) in sterile water was added to the

lagoon at 2 to 3 ml hour<sup>-1</sup> to obtain a final concentration of 1 mM for induced mutagenesis (Pharmacia Biotech Pump P-1). Samples were taken from the outflow of the lagoon, centrifuged, and supernatants were stored at 4°C. Samples were serially diluted before infecting TG1 cells with an appropriate reporter plasmid. Phage titers were analyzed by selection on ampicillin plates and colony counting. Single colonies were picked and grown overnight in 2× TY [ampicillin (100 µg ml<sup>-1</sup>)]. Phagemid DNA was purified using the QIAprep Spin Miniprep Kit (QIAGEN) and analyzed by sequencing (Eurofins Genomics). The DNA sequences of selected TFs are listed in fig. S11. For the *cl<sub>opt</sub>*/RFP enrichment assay, TG1 cells were infected with phage dilutions and plated on ampicillin plates. The ratio of white to red colonies was analyzed by colony counting, and white colonies were linked to *cl<sub>opt</sub>* infection by colony PCR.

### β-Lactamase reversion assay

TG1 cells were transformed with the reporter plasmid pLA230 and (i) *cloDF*-P<sub>BAD</sub>-*polA*-SpecR, (ii) *cloDF*-P<sub>BAD</sub>-EP<sub>polA</sub>-SpecR, (iii) *cloDF*-P<sub>BAD</sub>-MP4-SpecR, or (iv) *cloDF*-P<sub>BAD</sub>-MP6-SpecR and streaked out on LB plates supplemented with spectinomycin, kanamycin, and 1% (w/v) D-glucose. The plasmid pLA230 (21) carries a β-lactamase gene with the ochre stop codon TAA at amino acid position 26, which is located 230 bp downstream the origin of replication. The next day, single colonies were picked and grown to the mid-log phase in the presence of 1% D-glucose. Cultures were induced with 1% (w/v) arabinose and incubated for 24 hours at 37°C, 250 rpm (Stuart Shaking Incubator SI500). Cells were diluted and plated on LB plates in the presence or absence of ampicillin and 1% D-glucose and incubated overnight at 37°C. The ratio of the number of ampicillin-resistant colonies divided by the total number of colonies on LB plates was calculated.

### Rifampicin resistance assay

TG1 cells were transformed with MP4-SpecR or MP6-SpecR and plated on LB plates with the appropriate antibiotics and 1% (w/v) D-glucose. The next day, single colonies were picked and grown to the mid-log phase in the presence of 1% D-glucose. Next, cultures were induced with 1% (w/v) arabinose and incubated for 24 hours at 37°C, 250 rpm. Cells were diluted and plated on LB plates with 1% D-glucose in the presence or absence of rifampicin and incubated in the dark for 24 hours at 37°C. The ratio of the number of rifampicin-resistant colonies divided by the total number of colonies on LB plates was calculated.

### RFP mutation assay

S1030 cells carrying HP-ΔPS-ΔgIII-ΔgVI, pJPC12-ΔPS-P<sub>M,CS</sub>-RBS<sub>Bba\_B0034</sub>-gVI, and a mutagenesis plasmid (MP4-SpecR, MP6-SpecR, P<sub>Lac</sub>-MP6-SpecR) were grown in 2× TY medium supplemented with antibiotics and 1% D-glucose until the OD<sub>600</sub> reached 0.4 to 0.6. Cells were infected with an RFP-carrying phagemid (pLITMUS-P<sub>Bba\_R0010</sub>-RFP-P<sub>Bba\_J23106</sub>-gIII) at MOI 1 and cultured in the presence or absence of an inducer (1% arabinose or 1 mM IPTG) for 20 hours at 30°C, 250 rpm. Phage supernatants were harvested the following day, and experiments were performed for three rounds of evolution in batch mode. TG1 cells were infected with diluted phage supernatants after each round and streaked out on LB plates supplemented with ampicillin. Plates were incubated overnight at 37°C, and the following day, white and red colonies were counted and the ratio was calculated.

## Reporter assay

TG1 cells were transformed with a pJPC12-derived reporter plasmid and a phagemid and selected overnight on agar plates at 37°C. The next day, single colonies were picked for each biological replicate and grown 3 to 5 hours in 1 ml of 2× TY supplemented with chloramphenicol (10 µg ml<sup>-1</sup>) and carbenicillin (10 µg ml<sup>-1</sup>) at 37°C, 250 rpm. The cultures were diluted to OD<sub>600</sub> 0.01, and 150 µl was added to each well of a 96-well plate. The absorbance at 600 nm, green fluorescence (excitation: 485 nm, emission: 520 nm), and red fluorescence (excitation: 585 nm, emission: 625 nm) were measured every 10 min in a Tecan Infinite F200 PRO plate reader (37°C, shaking between readings) until the *E. coli* cells reached stationary phase. For data analysis, fluorescence readings in the late-exponential phase (OD<sub>600</sub> of 0.2 for pLITMUS, OD<sub>600</sub> of 0.9 for pLITMUS\*) were used. Both absorbance and fluorescence were background corrected. The fluorescence was then normalized for the number of cells by dividing by the absorbance. The average of three or four biological replicates and the corresponding SD were calculated for each sample.

## Analysis of gene circuits

Single colonies from TG1 cells carrying a pJPC12-derived reporter plasmid and a p15A-derived TF plasmid were grown for 3 to 4 hours in 2 ml of 2× TY supplemented with chloramphenicol (5 µg/ml) and carbenicillin (5 µg/ml). The cultures were diluted to OD<sub>600</sub> 0.01 in a total volume of 150 µl in each well of a 96-well plate. Arabinose (0.01, 0.1, 1, and 2%) and *N*-(β-ketocaproyl)-L-homoserine lactone (3OC6-HSL; 0.01, 0.1, 1, and 2 µM) were added where appropriate. The absorbance at 600 nm, green fluorescence (excitation: 485 nm, emission: 520 nm), and red fluorescence (excitation: 585 nm, emission: 625 nm) were measured every 10 min in a Tecan Infinite F200 PRO plate reader (37°C, shaking between readings) until the *E. coli* population reached stationary phase. For data analysis, fluorescence readings in the mid- to late-exponential phase were used. Both absorbance and fluorescence were background corrected. The fluorescence was then normalized for the number of cells by dividing by the absorbance. The average of three biological replicates and the corresponding SD were calculated for each sample.

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/24/eaba2728/DC1>

[View/request a protocol for this paper from Bio-protocol.](#)

## REFERENCES AND NOTES

1. M. Madan Babu, S. A. Teichmann, Evolution of transcription factors and the gene regulatory network in *Escherichia coli*. *Nucleic Acids Res.* **31**, 1234–1244 (2003).
2. E. Pérez-Rueda, J. Collado-Vides, L. Segovia, Phylogenetic distribution of DNA-binding transcription factors in bacteria and archaea. *Comput. Biol. Chem.* **28**, 341–350 (2004).
3. H. H. McAdams, B. Srinivasan, A. P. Arkin, The evolution of genetic regulatory systems in bacteria. *Nat. Rev. Genet.* **5**, 169–178 (2004).
4. S. A. F. T. van Hijum, M. H. Medema, O. P. Kuipers, Mechanisms and evolution of control logic in prokaryotic transcriptional regulation. *Microbiol. Mol. Biol. Rev.* **73**, 481–509 (2009).
5. D. J. Lee, S. D. Minchin, S. J. W. Busby, Activating transcription in bacteria. *Annu. Rev. Microbiol.* **66**, 125–152 (2012).
6. M. Ptashne, Principles of a switch. *Nat. Chem. Biol.* **7**, 484–487 (2011).
7. S. J. Eisenbeis, M. S. Nasoff, S. A. Noble, L. P. Bracco, D. R. Dodds, M. H. Caruthers, Altered Cro repressors from engineered mutagenesis of a synthetic cro gene. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1084–1088 (1985).
8. S. J. Lee, M. Shirakawa, H. Akutsu, Y. Kyogoku, M. Shiraiishi, K. Kitano, M. Shin, E. Ohtsuka, M. Ikehara, Base sequence-specific interactions of operator DNA fragments with the lambda-cro repressor coupled with changes in their conformations. *EMBO J.* **6**, 1129–1135 (1987).
9. A. Poteete, K. Hehir, R. Sauer, Bacteriophage P22 Cro protein: Sequence, purification, and properties. *Biochemistry* **14**, 251–256 (1986).
10. G. del Solar, A. M. Hernández-Arriaga, F. X. Gomis-Rüth, M. Coll, M. Espinosa, A genetically economical family of plasmid-encoded transcriptional repressors involved in control of plasmid copy number. *J. Bacteriol.* **184**, 4943–4951 (2002).
11. F. D. Bushman, M. Ptashne, Turning λ Cro into a transcriptional activator. *Cell* **54**, 191–197 (1988).
12. A. K. Brödel, A. Jaramillo, M. Isalan, Engineering orthogonal dual transcription factors for multi-input synthetic promoters. *Nat. Commun.* **7**, 13858 (2016).
13. A. A. K. Nielsen, B. S. Der, J. Shin, P. Vaidyanathan, V. Paralanov, E. A. Strychalski, D. Ross, D. Densmore, C. A. Voigt, Genetic circuit design automation. *Science* **352**, aac7341 (2016).
14. T. S. Moon, C. Lou, A. Tamsir, B. C. Stanton, C. A. Voigt, Genetic programs constructed from layered logic gates in single cells. *Nature* **491**, 249–253 (2012).
15. A. K. Brödel, A. Jaramillo, M. Isalan, Intracellular directed evolution of proteins from combinatorial libraries based on conditional phage replication. *Nat. Protoc.* **12**, 1830–1843 (2017).
16. K. M. Esvelt, J. C. Carlson, D. R. Liu, A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499–503 (2011).
17. A. H. Badran, D. R. Liu, In vivo continuous directed evolution. *Curr. Opin. Chem. Biol.* **24**, 1–10 (2015).
18. F. D. Bushman, C. Shang, M. Ptashne, A single glutamic acid residue plays a key role in the transcriptional activation function of lambda repressor. *Cell* **58**, 1163–1171 (1989).
19. J. Peterson, G. J. Phillips, New pSC101-derivative cloning vectors with elevated copy numbers. *Plasmid* **59**, 193–201 (2008).
20. J. M. Allen, D. M. Simcha, N. G. Ericson, D. L. Alexander, J. T. Marquette, B. P. Van Biber, C. J. Troll, R. Karchin, J. H. Bielas, L. A. Loeb, M. Camps, Roles of DNA polymerase I in leading and lagging-strand replication defined by a high-resolution mutation footprint of ColE1 plasmid replication. *Nucleic Acids Res.* **39**, 7020–7033 (2011).
21. M. Camps, J. Naukkarinen, B. P. Johnson, L. A. Loeb, Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9727–9732 (2003).
22. A. H. Badran, D. R. Liu, Development of potent in vivo mutagenesis plasmids with broad mutational spectra. *Nat. Commun.* **6**, 8425 (2015).
23. R. Lutz, H. Bujard, Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I<sub>1</sub>-I<sub>2</sub> regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997).
24. J. G. Kim, Y. Takeda, B. W. Matthews, W. F. Anderson, Kinetic studies on Cro repressor-operator DNA interaction. *J. Mol. Biol.* **196**, 149–158 (1987).
25. N. Benson, P. Sugiono, P. Youderian, DNA Sequence Determinants of lambda Repressor Binding in Vivo. *Genetics* **118**, 21–29 (1988).
26. D. A. Siegle, J. C. Hu, Gene expression from plasmids containing the araBAD promoter at saturating inducer concentrations represents mixed populations. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8168–8172 (1997).
27. J. Hasty, D. McMillen, J. J. Collins, Engineered gene circuits. *Nature* **420**, 224–230 (2002).
28. J. W. Kotula, S. J. Kerns, L. A. Shakel, L. Siraj, J. J. Collins, J. C. Way, P. A. Silver, Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4838–4843 (2014).
29. A. S. Khalil, J. J. Collins, Synthetic biology: Applications come of age. *Nat. Rev. Genet.* **11**, 367–379 (2010).
30. S. Atsumi, J. W. Little, A synthetic phage λ regulatory circuit. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19045–19050 (2006).
31. J. G. Omichinski, P. V. Pedone, G. Felsenfeld, A. M. Gronenborn, G. M. Clore, The solution structure of a specific GAGA factor–DNA complex reveals a modular binding mode. *Nat. Struct. Biol.* **4**, 122–132 (1997).
32. T. L. Smith, R. T. Sauer, Dual regulation of open-complex formation and promoter clearance by Arc explains a novel repressor to activator switch. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8868–8872 (1996).
33. E. Pérez-Rueda, J. Collado-Vides, The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12. *Nucleic Acids Res.* **28**, 1838–1847 (2000).
34. R. A. Albright, B. W. Matthews, How Cro and λ-repressor distinguish between operators: The structural basis underlying a genetic switch. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3431–3436 (1998).
35. M. Piskacek, M. Havelka, M. Rezacova, A. Knight, The 9aaTAD transactivation domains: From Gal4 to p53. *PLOS ONE* **11**, e0162842 (2016).
36. J. Carlson, A. H. Badran, D. A. Guggiana-Nilo, D. R. Liu, Negative selection and stringency modulation in phage-assisted continuous evolution. *Nat. Chem. Biol.* **10**, 216–222 (2014).
37. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison III, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
38. Y. Schaerli, A. Munteanu, M. Gili, J. Cotterell, J. Sharpe, M. Isalan, A unified design space of synthetic stripe-forming networks. *Nat. Commun.* **5**, 4905 (2014).



39. M. Isalan, C. Lemerle, K. Michalodimitrakis, C. Horn, P. Beltrao, E. Raineri, M. Garriga-Canut, L. Serrano, Evolvability and hierarchy in rewired bacterial gene networks. *Nature* **452**, 840–845 (2008).
40. D. Huang, W. Holtz, M. Maharbiz, A genetic bistable switch utilizing nonlinear protein degradation. *J. Biol. Eng.* **6**, 9 (2012).
41. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

**Acknowledgments:** We thank D. B. Holst Hansen for support. **Funding:** This research was supported by the European Commission grant FP7-ICT-2013-10 (no. 610730, EVOPROG) and the BBSRC grant EVO-ENGINE BB/P020615/1. A.J. is funded by FP7-KBBE (no. 613745, PROMYS), H2020 Marie Skłodowska-Curie (no. 642738, MetaRNA), and EPSRC-BBSRC (no. BB/M017982/1, WISB center). M.I. is funded by New Investigator award no WT102944 from the Wellcome Trust U.K and by the Volkswagen Foundation. **Author contributions:** Conceived and designed the

experiments: A.K.B., A.J., and M.I. Performed the experiments: A.K.B. Analyzed the data: A.K.B. and M.I. Contributed reagents/materials/analysis tools: R.R., A.J., and M.I. Wrote the manuscript: A.K.B. and M.I. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. A set of plasmids, along with maps and sequences, were deposited in Addgene (IDs are listed in table S3).

Submitted 19 November 2019

Accepted 14 April 2020

Published 10 June 2020

10.1126/sciadv.aba2728

**Citation:** A. K. Brödel, R. Rodrigues, A. Jaramillo, M. Isalan, Accelerated evolution of a minimal 63-amino acid dual transcription factor. *Sci. Adv.* **6**, eaba2728 (2020).

## Accelerated evolution of a minimal 63–amino acid dual transcription factor

Andreas K. Brödel, Rui Rodrigues, Alfonso Jaramillo and Mark Isalan

*Sci Adv* **6** (24), eaba2728.

DOI: 10.1126/sciadv.aba2728

### ARTICLE TOOLS

<http://advances.sciencemag.org/content/6/24/eaba2728>

### SUPPLEMENTARY MATERIALS

<http://advances.sciencemag.org/content/suppl/2020/06/08/6.24.eaba2728.DC1>

### REFERENCES

This article cites 41 articles, 11 of which you can access for free  
<http://advances.sciencemag.org/content/6/24/eaba2728#BIBL>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).